Development
Learning
Lab

Designing Randomized Controlled Trials

Tony Gallenstein, Catholic University of America

Objectives

My objective is that by the end of the talk, participants will...

1. have a solid grasp on the intuition behind randomized controlled trials (RCTs),
2. know the key design features of RCTs,
3. understand challenges to a good RCT and how to solve them.

Building Intuition

The Goal of Impact Evaluation

Estimate the Average Treatment Effect

Average Treatment Effect (ATE) =

The average impact/effect of your program on the target population.

Randomized Controlled Trial (RCT)

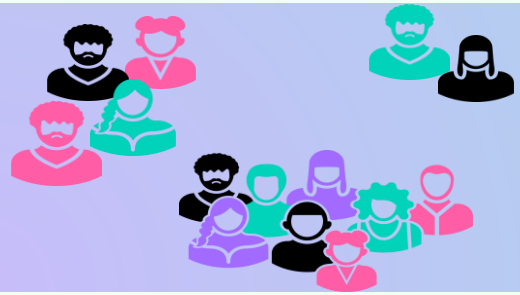
→ An RCT is a randomized impact evaluation.

→ Based on a couple simple principles:

1. Compare two groups: Treatment Group and Comparison Group
2. Assignment to the treatment and comparison groups is **Random.**

   ➔ the comparison group will be a **valid estimate of the counterfactual.**

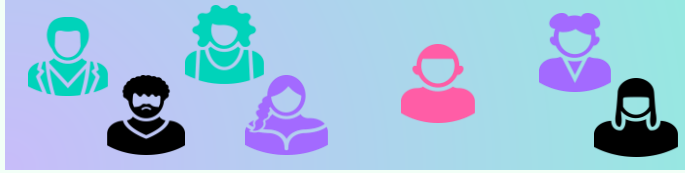
Randomly Assign to Treatment and Comparison Groups



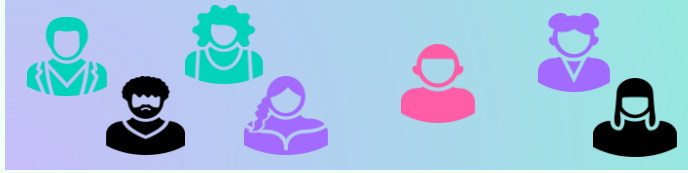
Sample Population

Randomly Assign to Treatment and Comparison Groups

Treatment Group



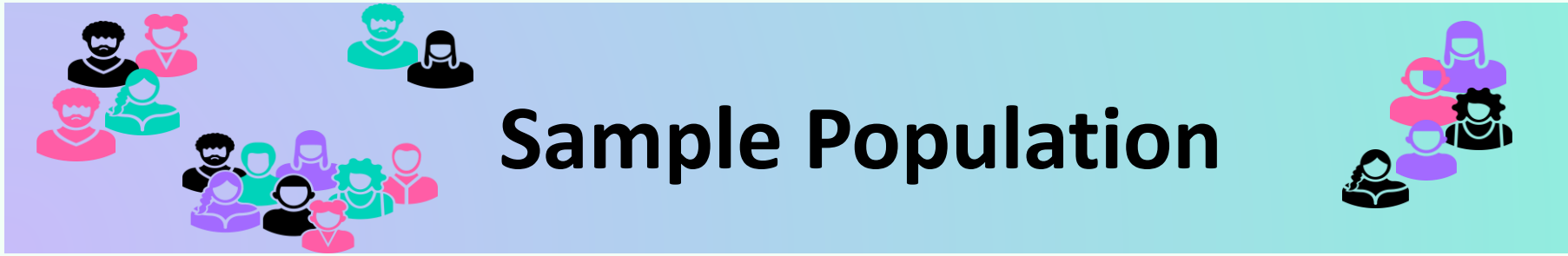
Comparison Group



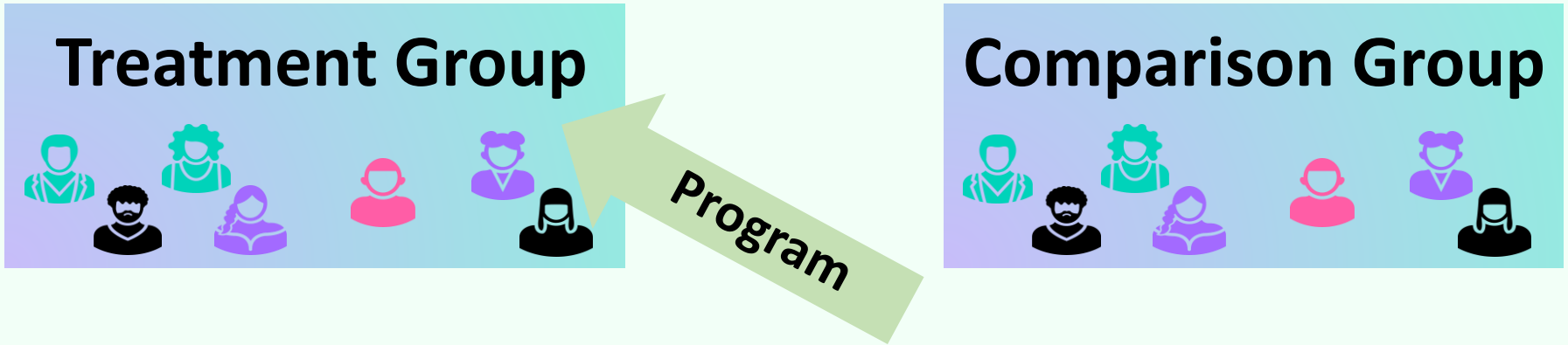
Without the Program

Average outcome for the Treatment Group = **Average outcome for the Comparison Group**

$$E[Y|P=1] = E[Y|P=0]$$



Randomly Assign to Treatment and Comparison Groups



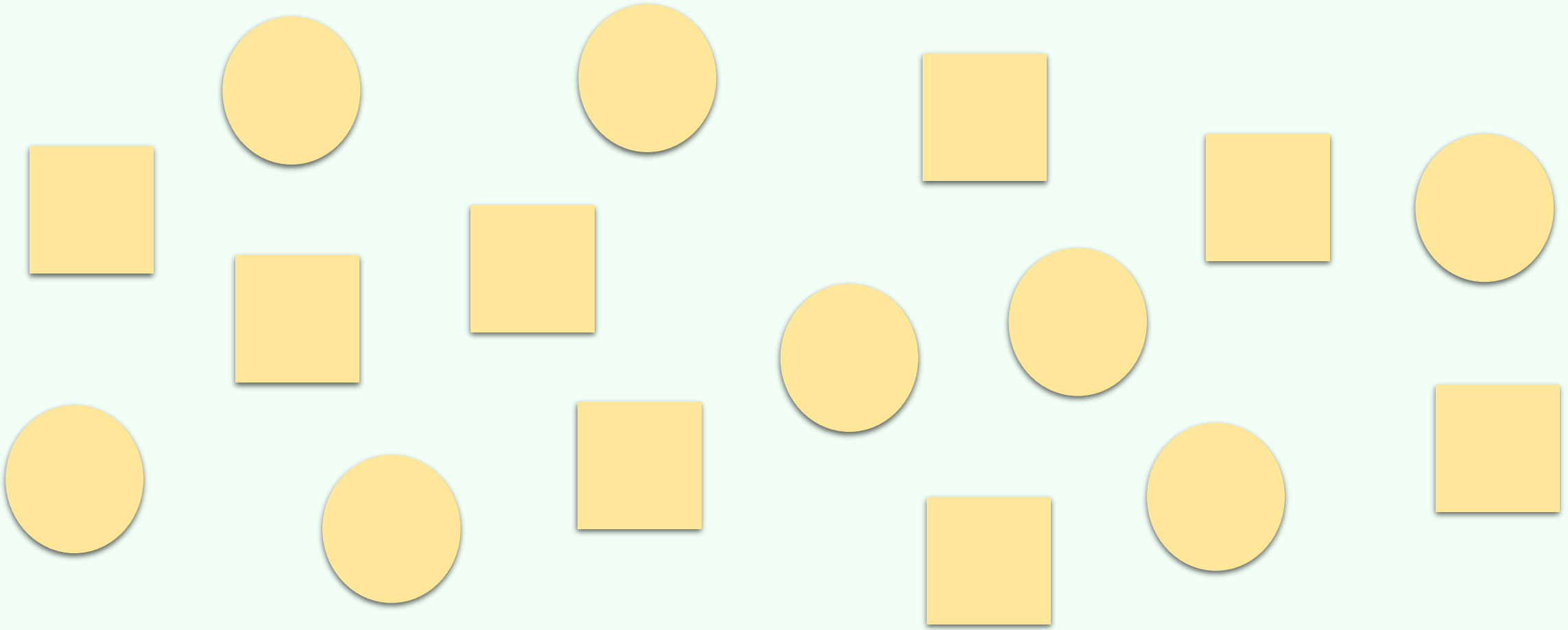
ATE = **Average outcome for the Treatment Group** − **Average outcome for the Comparison Group**

$$E[Y|P=1] - E[Y|P=0]$$

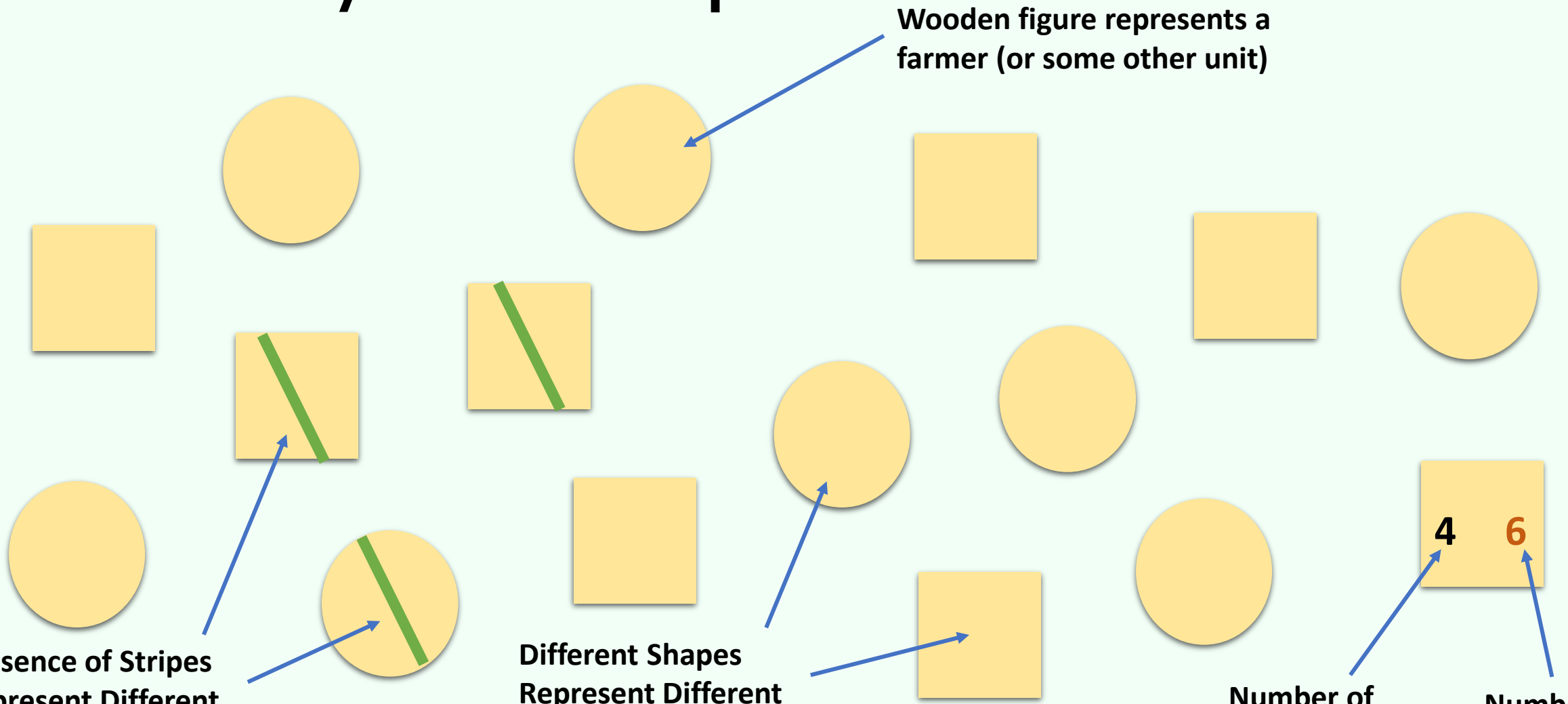
Stylized Case Study

Question	What is the impact of agricultural loans on farmers' crop yields?
Outcome	Bags of Maize per Acre (crop yields)
Program	Provide Loan

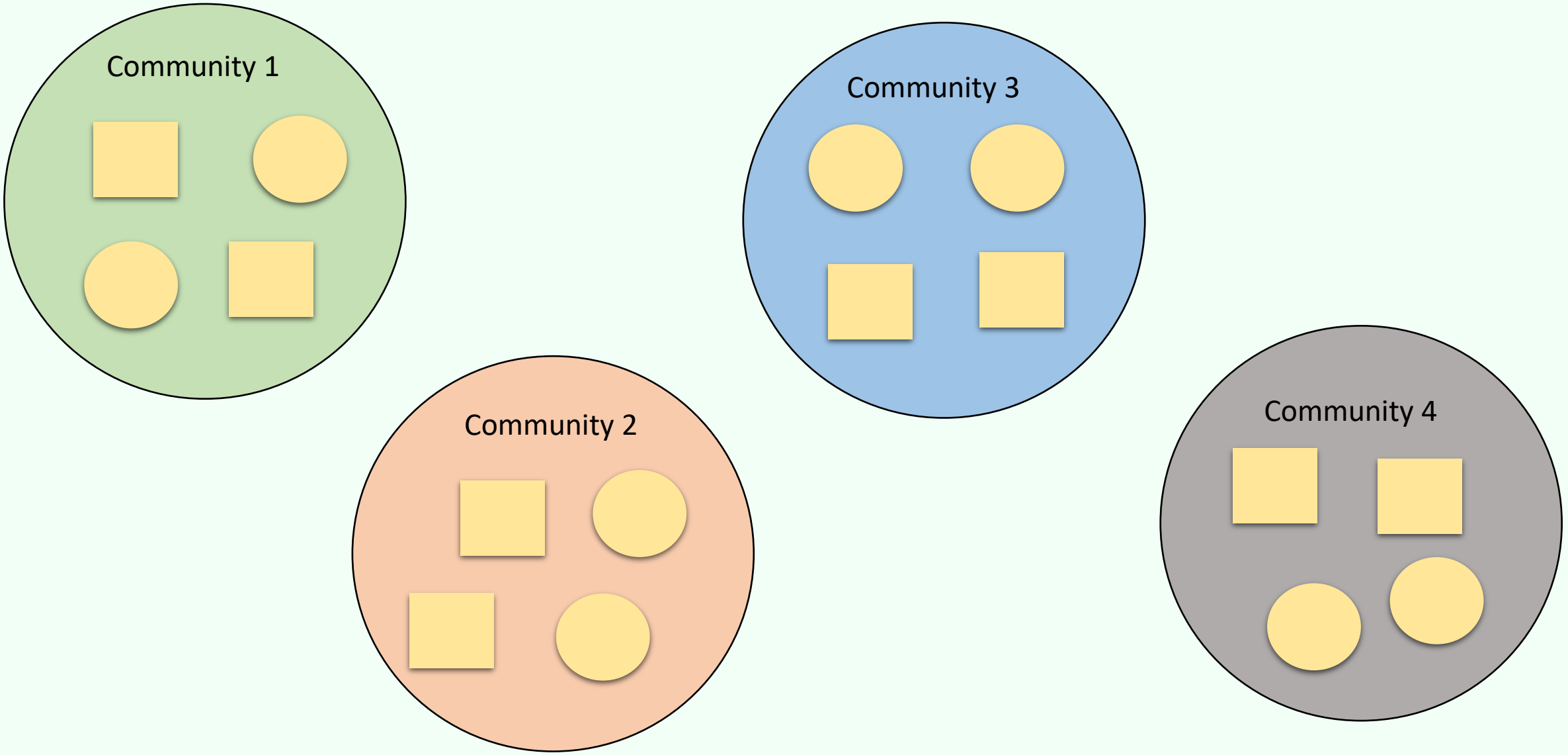
Our Stylized Sample



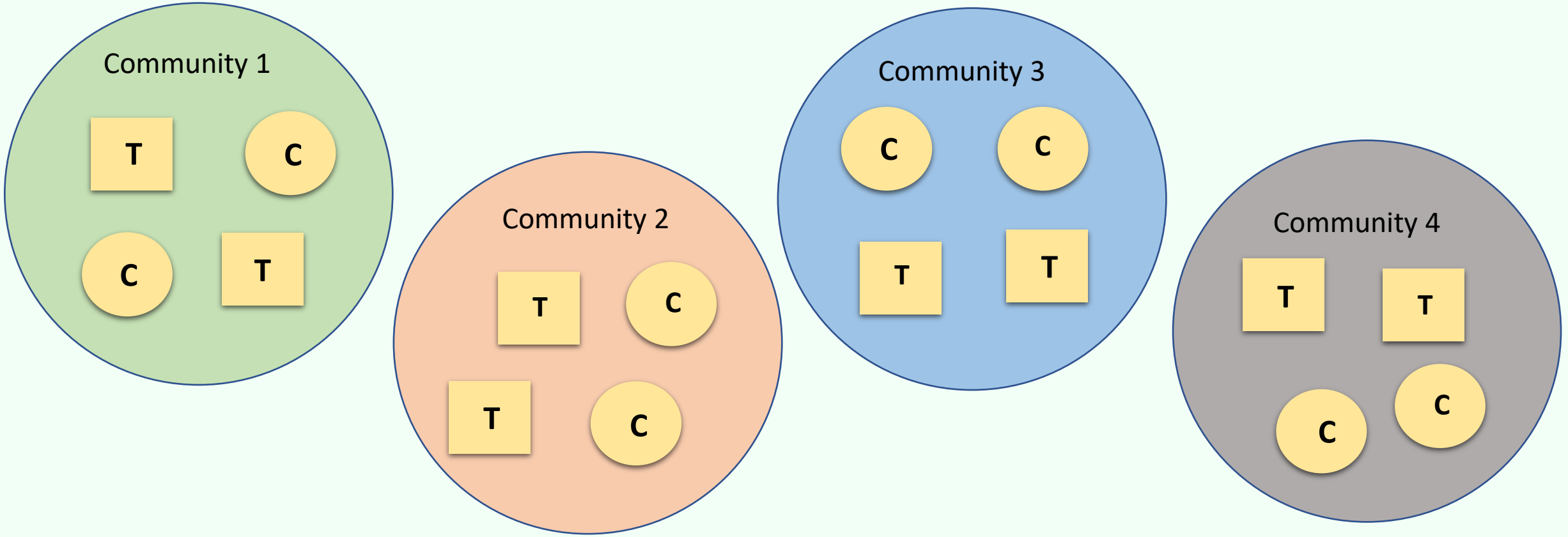
Our Stylized Sample



Our Stylized Sample

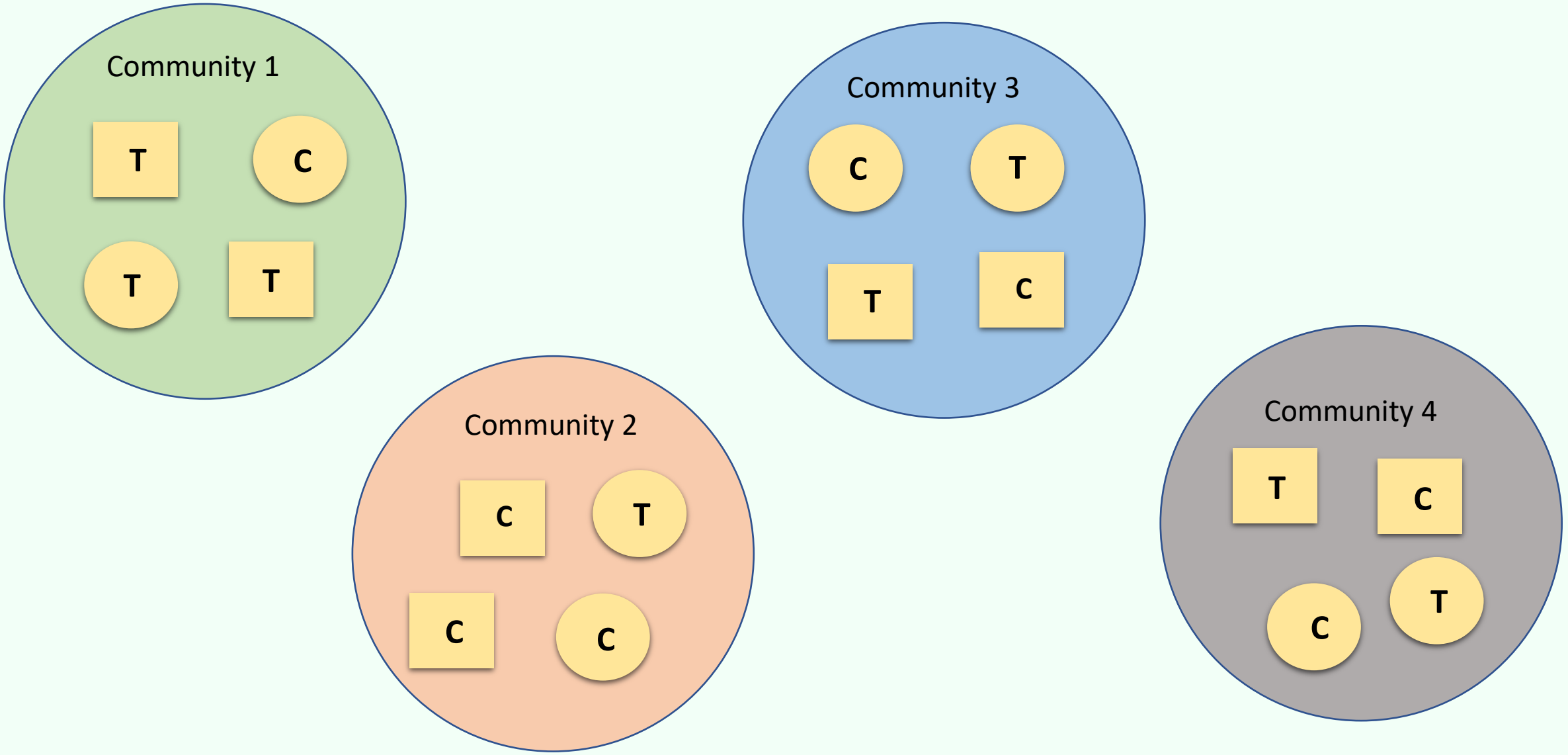


What is Wrong with This Design?

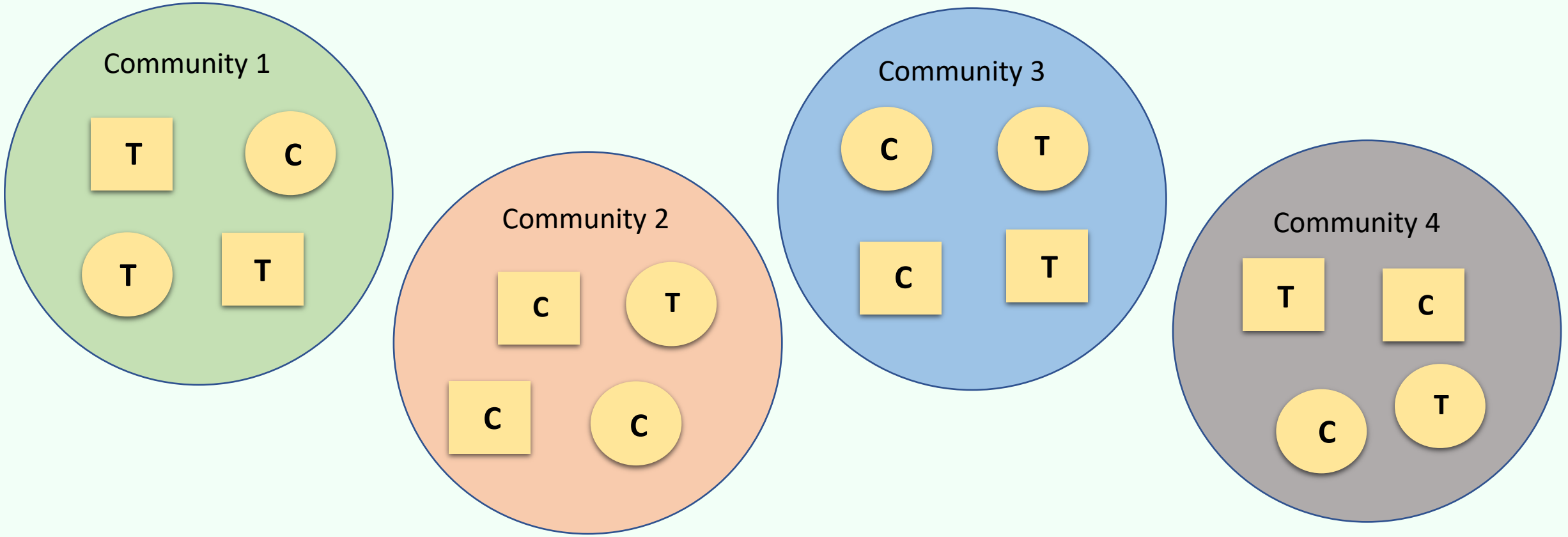


	$E[Y\|P=1]$	$E[Y\|P=0]$	Difference
With	7	3	4
Without	5	3	2

Randomization



Randomization



	$E[Y\|P=1]$	$E[Y\|P=0]$	Difference
With	6	4	2 $\widehat{ATE}$
Without	4	4	0

RCT Design

Choose a Level of Randomization

Level of Randomization = The unit of observation that is randomized.

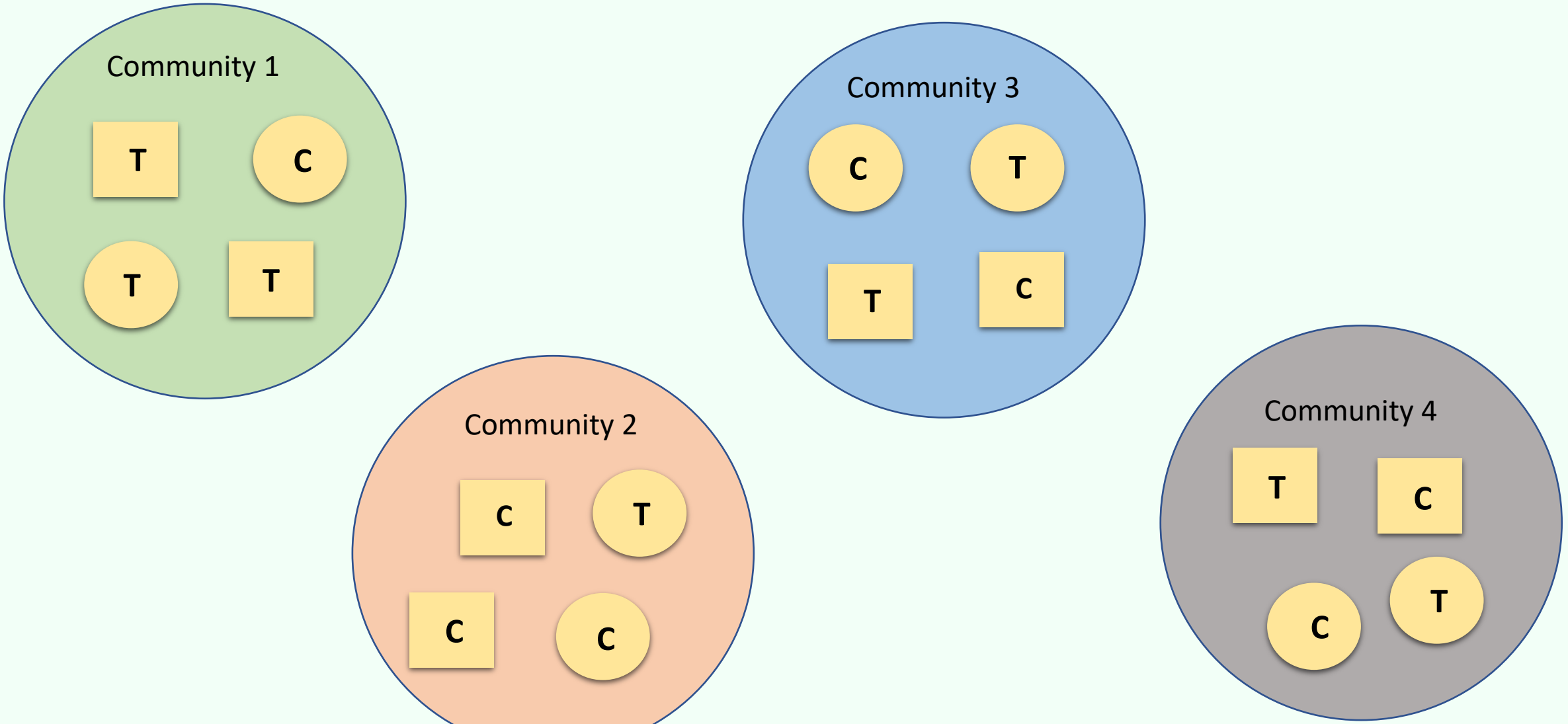
We need to choose the level of randomization:

Household level, OR

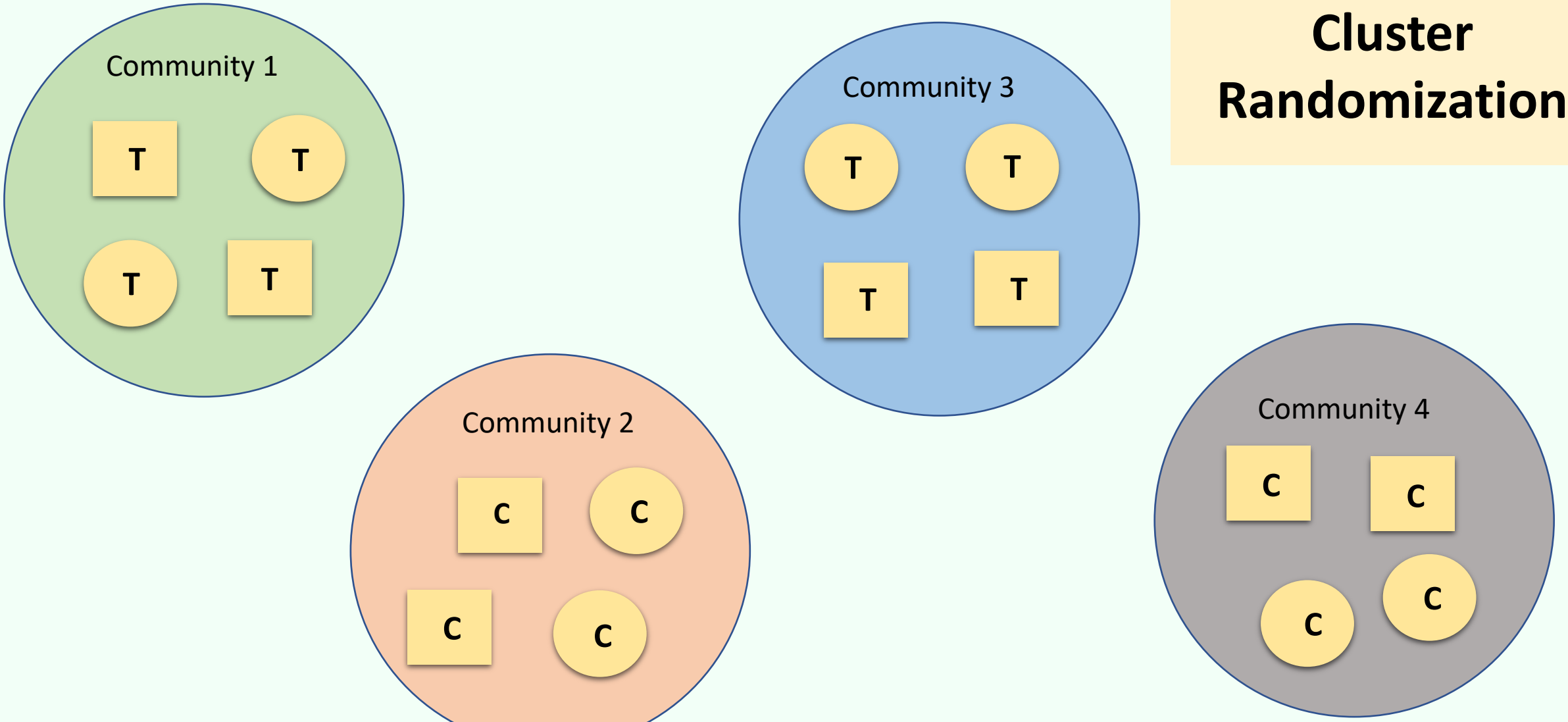
Community level

Higher level randomization is typically called a “cluster” randomization.

Farmer Level Randomization



Community Level Randomization



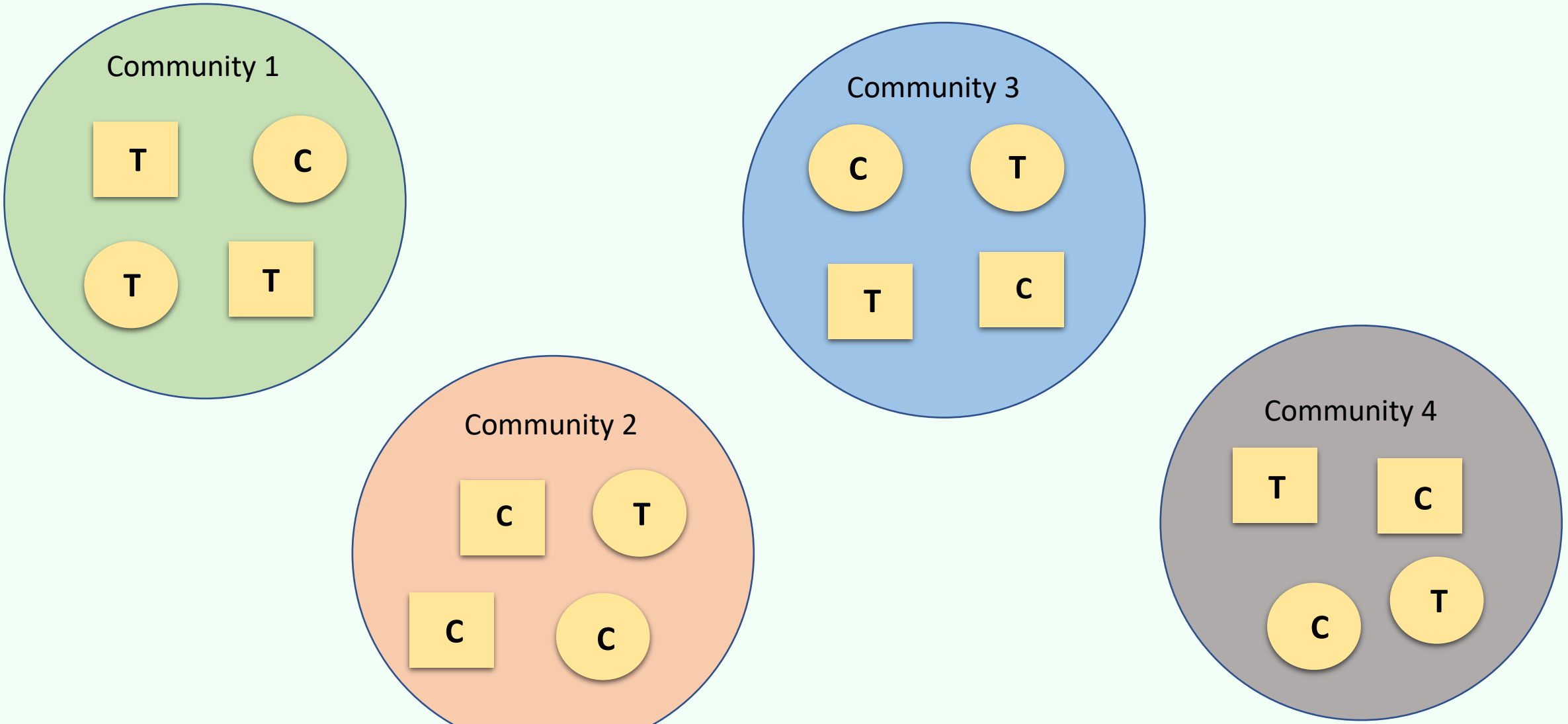
Cluster Randomization

Choose to Stratify or Not

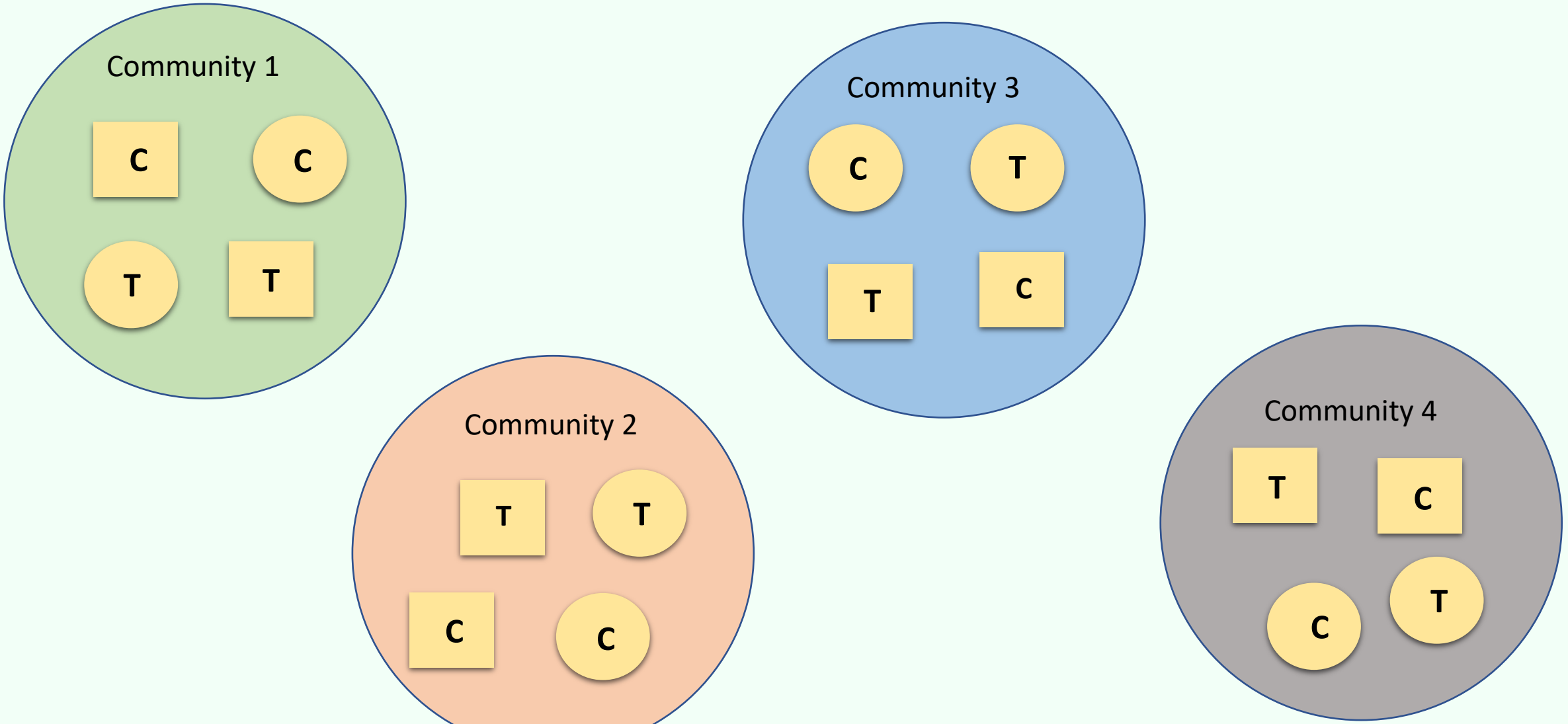
Stratification is when we randomly assign the treatment within subgroups of the total sample population.

→ This improves comparability between treatment and comparison groups

Farmer Level Randomization



Farmer Level Randomization / Stratify by Community



Randomization Method

How you implement the randomization depends on the context.

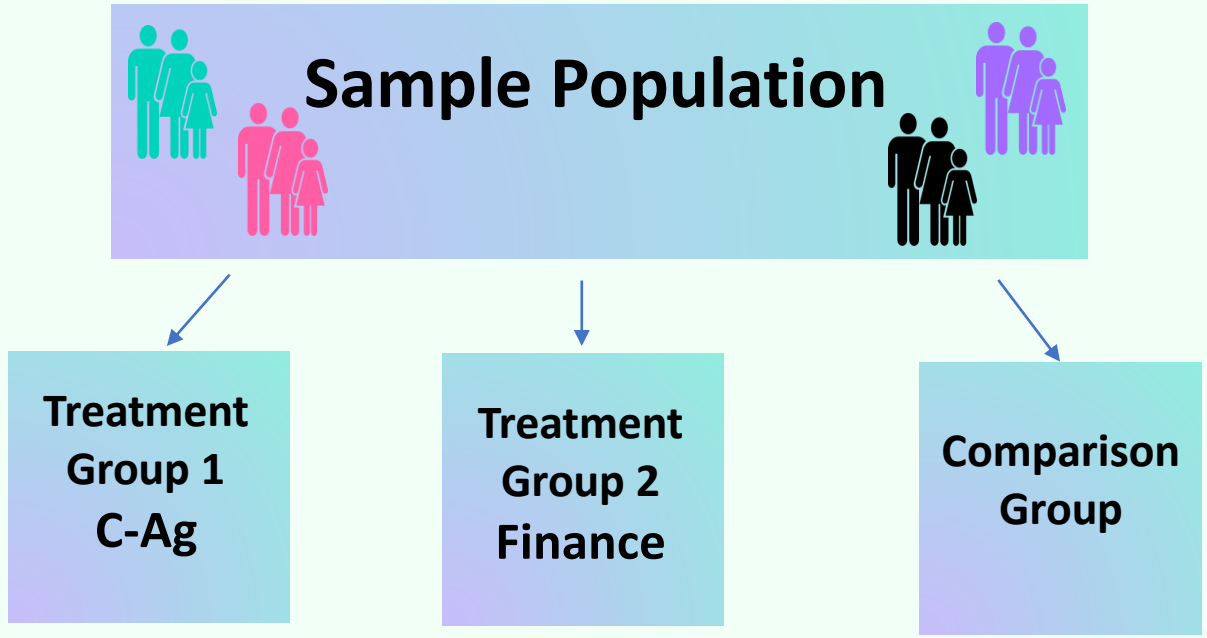
Randomization Methods: Multiple Treatments

When?

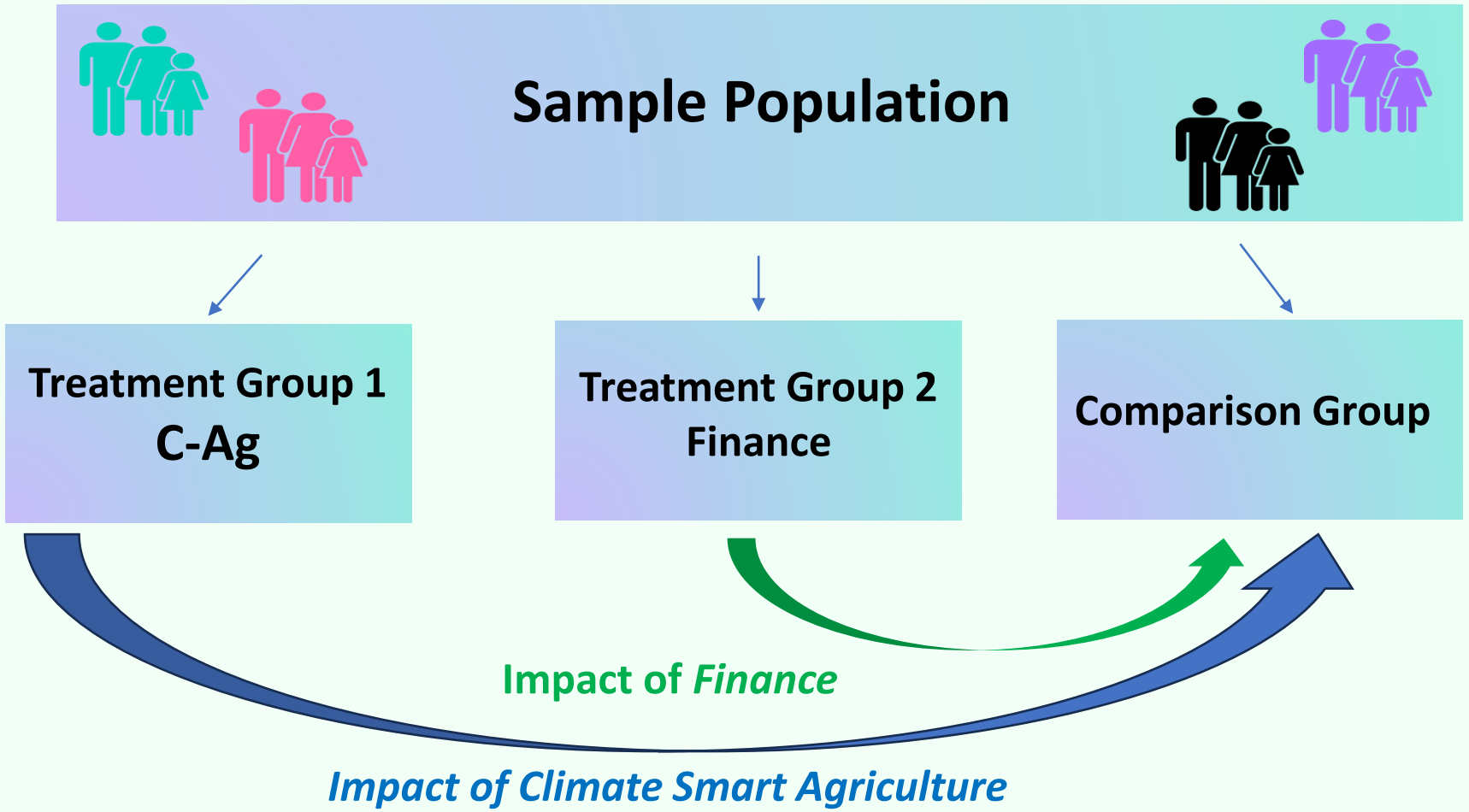
- multiple programs
- program has multiple parts
  - Example: Training on C-Ag + Finance

How?

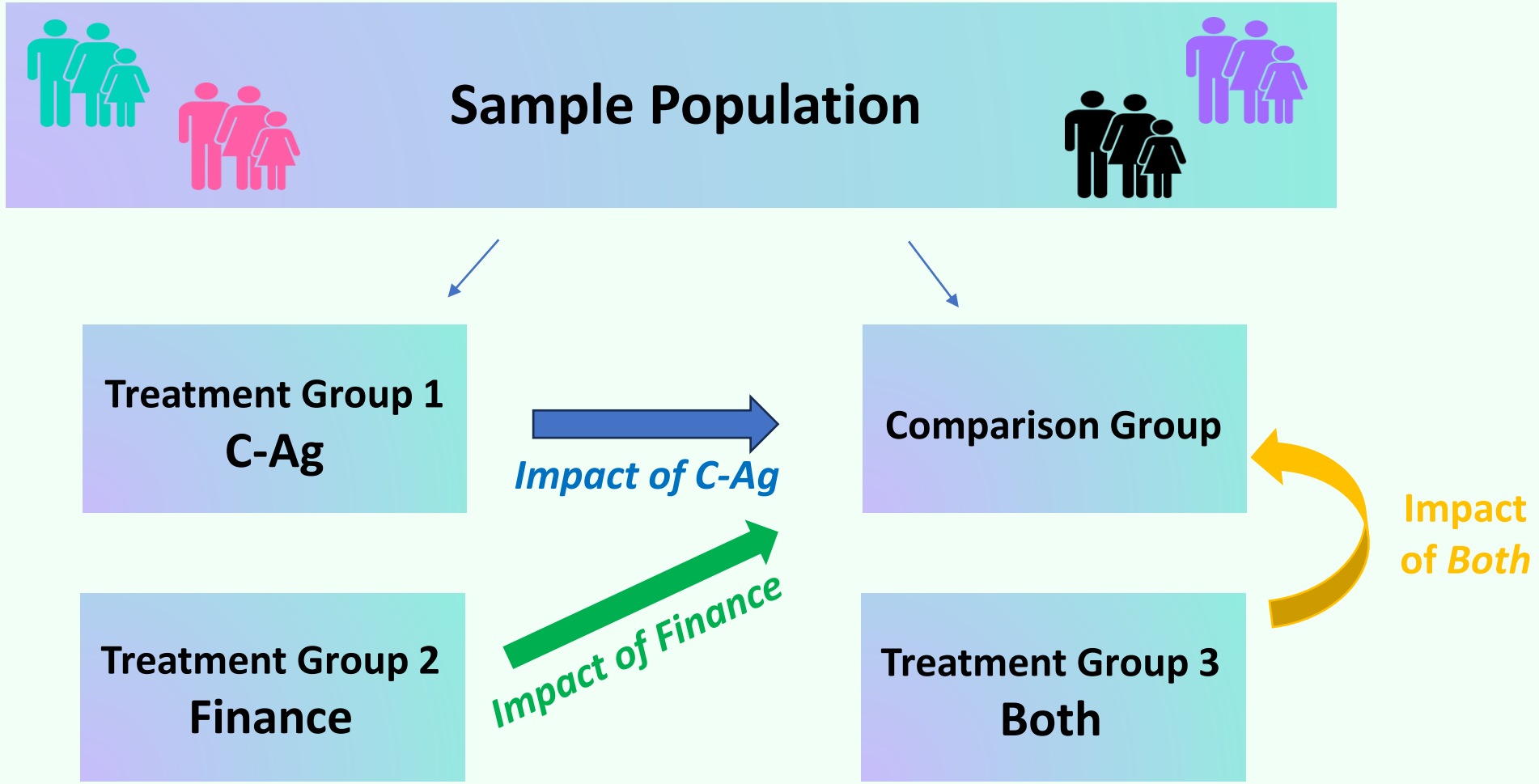
- Randomly assign sample to multiple treatment groups and one comparison group.



Randomization Methods: Multiple Treatments



Randomization Methods: Multiple Treatments



Randomization Methods: Lottery

When?

- Recipients apply for the program.

How?

- Of the applicants, use a random lottery to determine who receives (treatment group) and who does not (comparison group).

Randomization Methods: Phase-in

When?

- The entire sample will eventually receive the program.
- Funding or logistic constraints dictate that not all the sample will receive the program in the beginning.

How?

- Randomly assign who will receive the program first (treatment group) and who will receive it later (comparison group).

Randomization Methods: Encouragement Design

When?

- A program is available to the entire sample / you cannot exclude the comparison group.

How?

- Randomly assign the sample to receive an encouragement to participate in the program.
- (Need high participation rate to be effective).

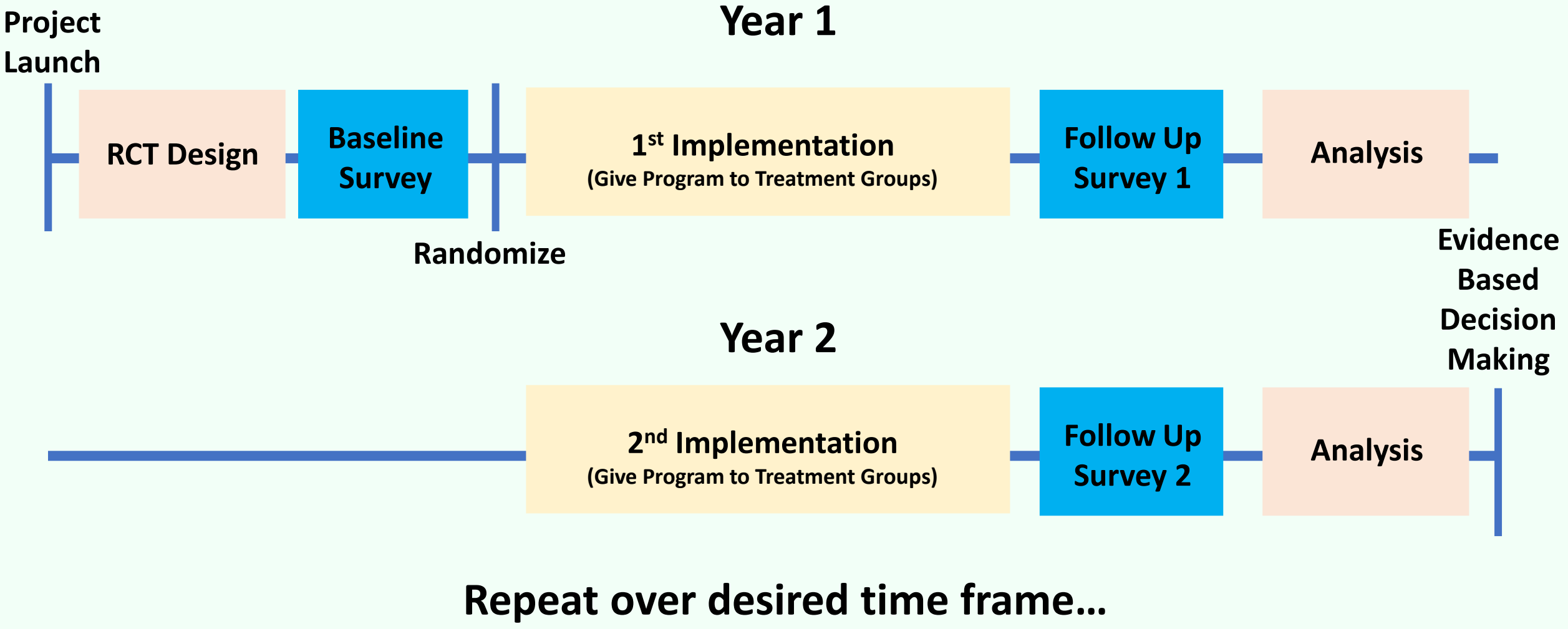
Now you are Ready to Randomize

Based on your design, conduct the randomization.

Timeline


Project Launch
RCT Design
Baseline Survey
Randomize
Implement Program
(Give Program to Treatment Groups)
Follow Up Survey
Analysis
Evidence Based Decision Making

Timeline – Long Run Impacts



Repeat over desired time frame…

Challenges

Imperfect Compliance

Imperfect compliance occurs when

- Not everyone assigned to the treatment group receives the program.
- Some comparison individuals somehow do receive the program.

This is particularly likely when a program is voluntary.

Imperfect Compliance

- When there is Imperfect compliance we estimate the **"Intent to Treat Effect"** or **ITT**.

Where the T indicates random assignment to the treatment group

$$\widehat{ITT} = E\left[Y_i^1 \middle| T_i = 1\right] - E\left[Y_i^0 \middle| T_i = 0\right]$$

- Sometimes we prefer to estimate the ITT to the ATE for programs that are inherently voluntary. (We estimate ITT in encouragement design RCTs)

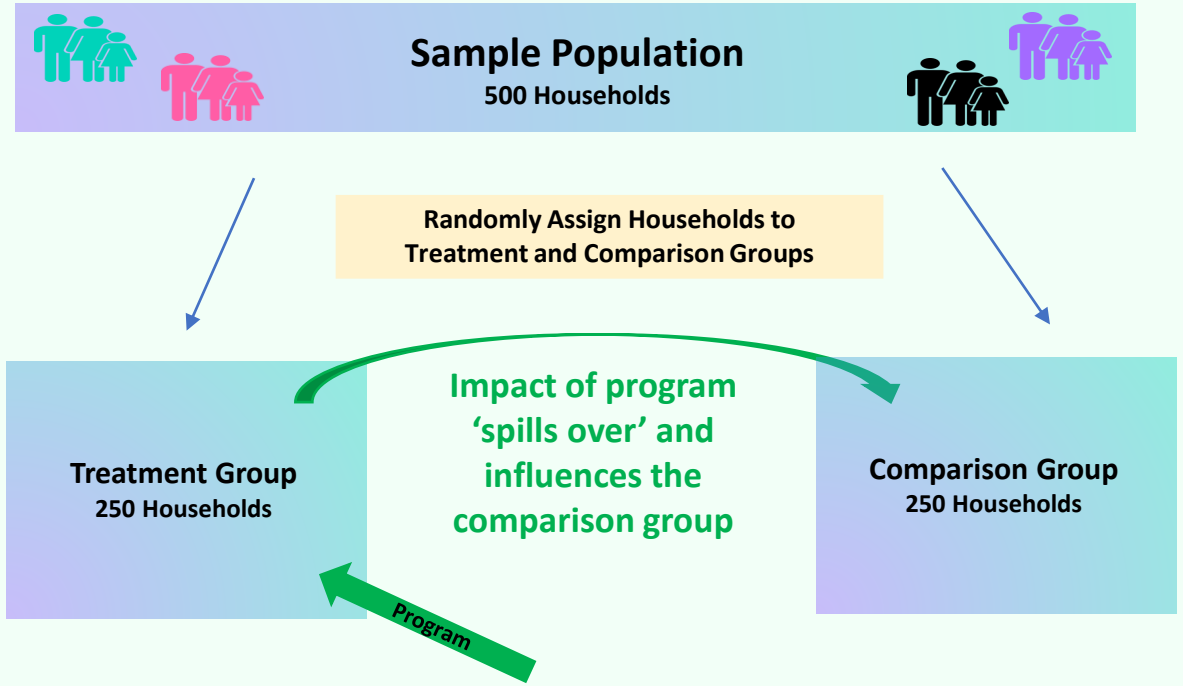
	$E[Y\|P=1]$	$E[Y\|P=0]$	Difference
With	5.5	4	1.5 $\widehat{ITT}$

Sources of Bias

1. Spillover effects
2. Attrition
3. Behavioral Bias
4. Equilibrium Effects

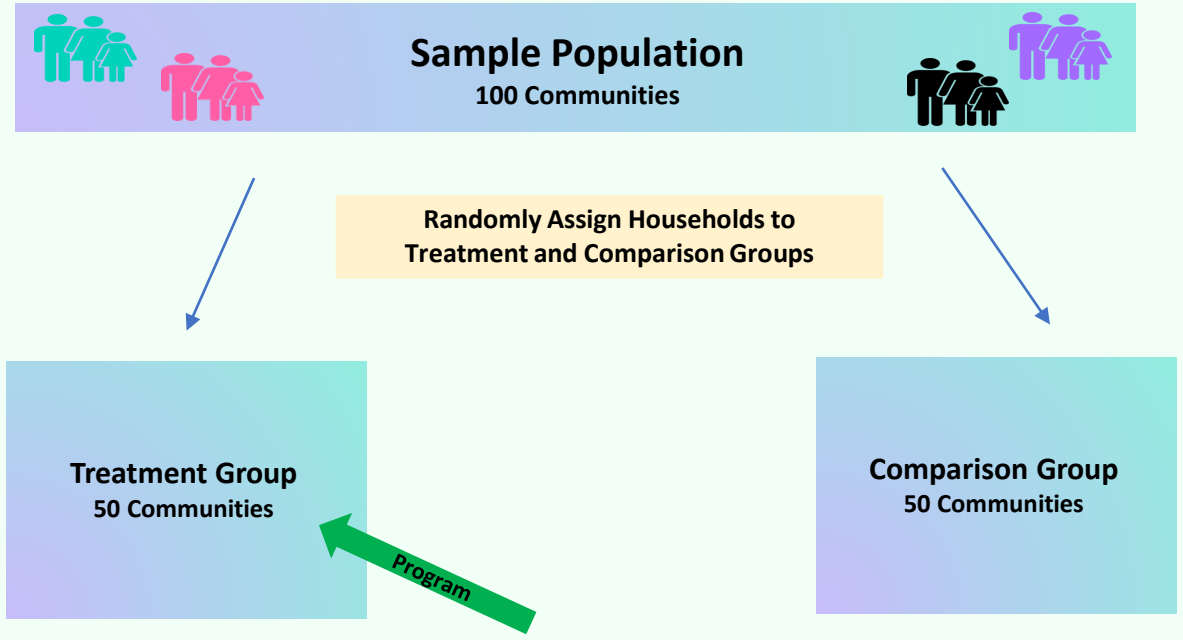
Spill Overs

- Spillover = The treatment has an impact on the comparison group.
- Problem: Biased estimate of the treatment effect.



Spill Overs

- Spillover = The treatment has an impact on the comparison group.
- Problem: Biased estimate of the treatment effect.
- Solution: Randomize at a higher level (cluster randomize).



Attrition

Attrition = Occurs when there is a failure to acquire data on individuals in the sample population.

Problem:

- If attritors are systematically different than the sample ➔ Biased Estimate of ATE.

Solutions:

- Ex Ante:
  - Track entire sample population in the survey, even if they left the program.

Behavioral Bias

- Hawthorne Effects: When the treatment group is impacted by being assigned to the treatment group ➔ Biased Estimate of ATE.
  - *Participants may work harder to ensure a pilot program is a success*

- John Henry Effects: When the comparison group is impacted by being assigned to the comparison group ➔ Biased Estimate of ATE.
  - *Individuals work harder to compensate for not having the treatment.*

- Solution:

1. Do not reveal that individuals are in a “treatment” or “comparison” group. (sometimes this can be unethical).
2. Do not reveal what outcomes you are focusing on.

Equilibrium Effects

Program has a different impact when given to everyone rather than a subsample.

Solutions:

→ Experimentally vary the intensity of implementation.

Development Learning Lab

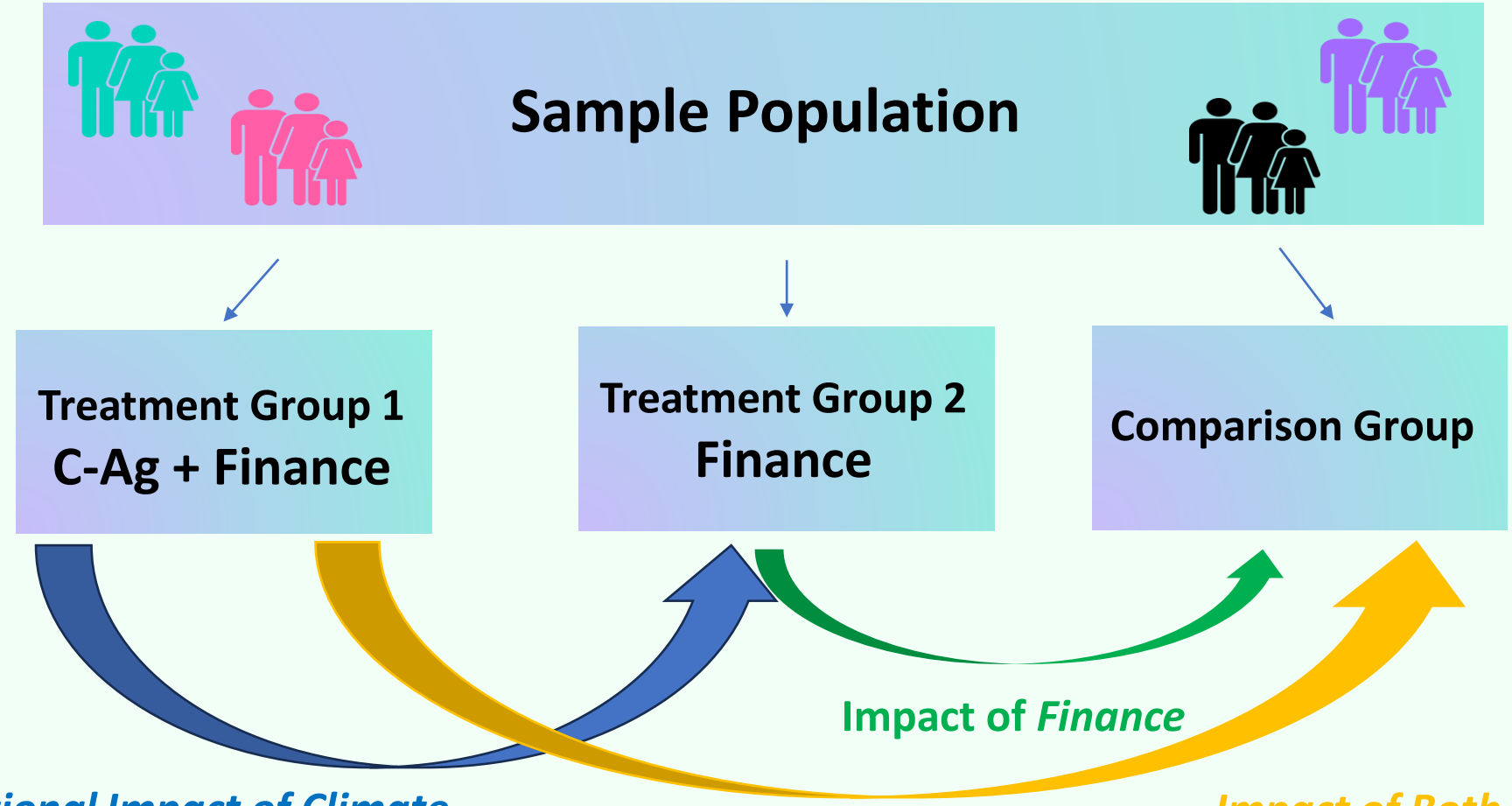
CMI CHR. MICHELSEN INSTITUTE

UNIVERSITY OF BERGEN

NHH

SNF

Randomization Methods: Multiple Treatments



Validate the Randomization

- Randomly assigning the treatment status should create comparability.
- But sometimes it does not (example: sample size too small).

- To Validate the Randomization: Use a **Balancing Test**
  - Compare average characteristics between the treatment and comparison groups to ensure they are the same on average in both groups.
  - Need a Baseline Survey (Survey before the program is implemented).

Discussion Questions

1. At what level is your program operating?
    1. Individual level? Household level? Community-wide?
2. At what level do you wish to capture the program's impact?
    1. Think about this in relation to each of your outcomes of interest.
3. If you were to randomize your program, how would you do it in practice?
    1. What would the treatments be?
    2. Level of randomization?
    3. Will you stratify?
    4. Which randomization method would you use?
    5. Will you have perfect or imperfect compliance?
    6. How would you limit spillovers/attrition/behavioral biases?
    7. Would the control group receive anything?
    8. What is your evaluation timeline?